COMPARISON OF STOCHASTIC RADIAL BASIS FUNCTION AND PEST FOR AUTOMATIC CALIBRATION OF COMPUTATIONALLY EXPENSIVE GROUNDWATER MODELS WITH APPLICATION TO MIYUN-HUAI-SHUN AQUIFER

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Ying Wan

January 2013

ABSTRACT

Groundwater numerical models have been widely used as effective tools to analyze and manage water resources. However, the accuracy and reliability of a groundwater numerical model largely depends on model parameters calibration, which is extremely computationally expensive. Therefore, it is highly desirable that efficient optimization algorithms be applied to automatic calibration problems. In this study, we compare the performance of three optimization algorithms and propose a new hybrid method. The algorithms are applied to calibration of a model for part of Beijing water supply.

We first outline the three algorithms and briefly describe our hybrid method. The first algorithm referred as PEST in this paper is the Gauss-Marquardt-Levenberg (GML) method including truncated singular value decomposition, which is widely applied in the field of model parameter calibration. As the second one, CMAES_P is a "PEST compatible" implementation of CMA-ES (Covariance Matrix Adaptation Evolution Strategy) global optimization algorithm. PEST derivative-based algorithm and CMAES_P are both encapsulated in the automated parameter optimization software PEST, which has advanced predictive analysis and regularization features to minimize user-specified objective functions. The third one, called Stochastic Radial Basis Function (Stochastic RBF) method, is developed by Regis and Shoemaker (2007), which utilizes radial basis function as the response surface model to approximate the expensive objective function. Our new hybrid method combines Stochastic RBF and PEST derivative-based algorithm, which provides PEST derivative-based algorithm with the starting points found by Stochastic RBF.

This paper compares the performances of the aforementioned four algorithms for automatic parameter calibration of a groundwater model on three 28-parameter cases and two synthetic test function calibration problems. We employ the following characteristics as our comparison criteria on all the cases: (1) efficiency in giving good objective function for a given number of function evaluations; (2) performance for different statistical criteria; (3) variability of solutions in multiple trials; (4) improvements if more function evaluations are performed. On the basis of 20 trials, the results indicate that Stochastic RBF is best among the three and CMAES_P is superior to PEST. In addition, our hybrid method still failed to beat Stochastic RBF in highly computationally expensive nonlinear cases.

To sum up, our results show that Stochastic RBF method is a more efficient alternative to PEST for automatic parameter calibration of computationally expensive groundwater models.

BIOGRAPHICAL SKETCH

Ying Wan was born in November 1987 in Hunan, China. After graduating from high school in 2005 in Hunan, China, she joined the Wuhan University, one of the top universities in China. She obtained a Bachelor of Science in Hydrology and Water Resources System at Wuhan University in June 2009. In 2009, she started her studies in Civil & Environmental Engineering MS/PhD program at Cornell University with Prof. Christine Shoemaker as an advisor. Her area of research includes model inversion/calibration and multi-objective optimization for numerical models applied to groundwater flow and transport.

This thesis is dedicated to my family and friends.

ACKNOWLEDGMENTS

Thanks to my advisor Prof. Christine Shoemaker for her guidance, her time and financial support. Through her passion for using Mathematics to solve real world environmental problems, she has been a key to my development in optimization linked to groundwater flow and transport problems.

Thanks to the CEE department, for years of funding as a teaching assistant.

Thanks to my committee members, who provided occasional and exceptionally helpful advice.

Thanks to Amandeep Singh, Suenee Tan and Yilun Wang from our EWRS (Environmental Water Resources and Systems) group for their kindly help and support.

Finally, thanks to my family for their constant and unwavering moral support, as well as the core values of hard work and resilience they have transmitted to me and without which this work could not have been done.

TABLE OF CONTENTS

BIOGRAPHICAL SKETCH III

DEDICATION IV

ACKNOWLEDGEMENTS V

TABLE OF CONTENTS VI

LIST OF FIGURES VIII

LIST OF TABLES IX

1 INTRODUCTION 1

2 THE PROBLEM OF MODEL CALIBRATION 9

2.1 GROUNDWATER FLOW MODELING 9

2.2 FORMULATION FOR OPTIMIZATION 10

3 OPTIMIZATION ALGORITHM DESCRIPTION 15

3.1 PEST DERIVATIVE-BASED ALGORITHM 15

3.2 CMAES_P 16

3.3 STOCHASTIC RBF 17

4 APPLICATION OF AUTOMATED CALIBRATION TO GROUNDWATER NUMERICAL MODEL 20

4.1 GENERAL DESCRIPTION OF STUDY AQUIFER 20

4.2 CALIBRATION AND STUDY CASES 22

4.2.1 Parameterization 22

4.2.2 Study Cases 23

4.2.3 Hypothetical Illustrative Example 25

4.3 NUMERICAL RESULTS 29

4.3.1 Analysis of Average Results for Multiple Cases 31

4.3.2 Model Simulation Results 33
4.3.3 Variability in Multiple Trials 35
4.3.4 Results for More Iterations 39
4.3.5 Combination of Stochastic RBF and PEST 41
4.3.6 Multiple Performance Criteria 44
5 APPLICATION OF AUTOMATED CALIBRATION TO TEST FUNCTIONS 48
5.1 TEST FUNCTION PROBLEMS SETUP 48
5.2 RESULTS 50
6 DISCUSSION AND CONCLUSIONS 52
BIBLIOGRAPHY 54

LIST OF FIGURES

4.1 Miyun-Huai-Shun aquifer site and surroundings. Location with hydraulic head observation wells of the study area, zones of model for parameters [*Wei*, 2001]21

4.2 Relationship between Hydraulic Conductivity and Specific Yield according to Previous Research from [*Wang et al.*, 2004] when used in Construction of Case 3....29

4.3 Best objective function value for each algorithm for 300 evaluations averaged over 20 trials: (a) Case 1: Fine Grid model for MHSA. (b) Case 2: Coarse Grid model for MHSA. (c) Case 3: Hypothetical Case. The lowest curve is the best. (MHSA is the model for the Miyun-Huai-Shun Aquifer).....32

4.4 Simulated hydraulic head vs observed hydraulic head from fine grid Beijing Groundwater model with parameter calibrated by (a) Stochastic RBF, (b) PEST and (c) CMAES_P respectively, after 300 simulations (calibration period + validation period).....34

4.5 Box plot of best solution for each algorithm based on 20 trials after 300 function evaluations: (a) Case 1: Fine Grid model for MHSA. (b) Case 2: Coarse Grid model for MHSA. (c) Case 3: Hypothetical Case. Smaller box means more consistency across all trials. (MHSA is the model for the Miyun-Huai-Shun Aquifer) 38

4.6 Average best objective function value for each algorithm formed as function of number of function evaluations of $f(h(\mathrm{x}))$ over 20 trials up to 500 simulations: (a) Fine Grid Case. (b) Coarse Grid Case. (c) Hypothetical Case. 40

4.7 Comparison of algorithm performance including method of combining Stochastic RBF and PEST (start PEST with average parameters from 100^{th} iteration of Stochastic RBF over 20 trials): (a) Case 1: Fine Grid model for MHSA. (b) Case 2: Coarse Grid model for MHSA. (c) Case 3: Hypothetical Case. The lowest curve is the best. (MHSA is the model for the Miyun-Huai-Shun Aquifer) 43

LIST OF TABLES

4.1 Range of Calibrated Parameters for the Miyun-Huai-Shun Groundwater Aquifer.....25

4.2 Values for Parameters of Hydraulic Conductivity and Specific Yield from Previous Research of Analyzed results of Well Test [*Wang et al.*, 2004] 27

4.3 True Parameter Values of Hypothetical Case (Case 3).....28

4.4 Average and Standard Deviation of the Objective Function Value Produced by Each Algorithm and Each Case after 300 Function Evaluations based on 20 trials.....37

4.5 Average and Standard Deviation of the Objective Function Value Produced by Each Algorithm and Each Case after 500 Function Evaluations based on 20 trials.....41

4.6 Multiple Performance Criteria results for all three Cases as well as all algorithms 47

CHAPTER 1

INTRODUCTION

Accurate and efficient calibration of parameters in groundwater models is important to improve the model predictive ability. Efficiency is important because groundwater models involve the solution of partial differential equations over grids with thousands or millions of node, so each simulation for a set of trial values of parameter vectors can take many minutes or hours even when run on parallel nodes. For this reason, we need accurate optimization algorithms for calibration of groundwater models that do not require a large number of simulations to get a good answer.

The most widely used method for groundwater calibration is PEST [*Doherty*, 2010]. The core of PEST is the Levenberg-Marquardt (LM) algorithm, which is a derivative-based local optimization method that will find a local minimum. Doherty and co-workers have made series of improvements to PEST [*Doherty and Johnston*, 2003; *Doherty and Welter*, 2010; *Doherty*, 2012] and the authors have offered many training sessions, which have contributed to the propagation of its use. There are also other similar optimization methods directed at groundwater calibration including UCODE [*Poeter et al.*, 2005; *Hill and Tiedeman*, 2007] which is set up for use with MODFLOW, and iTOUGH2 [*Finsterle*, 2007] which is designed to work with TOUGH2 [*Pruess et al.*, 1999]. All of the methods PEST, iTough and UCODE contain the Levenberg-Marquardt algorithm and most of the improvements (e.g. the

automatic user intervention) mentioned into PEST were also later incorporated into UCODE and iTough. *Finsterle and Zhang* [2011] has done preliminary comparison of iTough to PEST on a TOUGH2 code and found the two algorithms performed similarly.

However, there are problems associated with using a derivative-based local optimization algorithm for groundwater remediation problem. Effort to calculate accurate derivatives and the existence of multiple local minimum are the two problems that can arise with use of the local optimization algorithm. For complex groundwater models, derivatives can be difficult or impossible to obtain analytically (even with automatic differentiation). So for calibration purpose, PEST and the other derivative-based methods usually calculate derivatives by finite difference method ($(h(x_1,\mathrm{K}\ ,x_i,\mathrm{K}\ ,x_N)-h(x_1,\mathrm{K}\ ,x_i+\Delta x_i,\mathrm{K}\ ,x_N))/\Delta x_i, i=1,2,\mathrm{K}\ N$). This means that for an N dimensional parameter calibration problem, one needs to do N groundwater model simulations in each iteration. This requires many groundwater model simulations, which is what we are trying to avoid [*Doherty*, 2010]. Automatic User Intervention method is designed to reduce the number of finite difference calculation by ignoring some of the partial derivatives and we use this feature in the version of PEST against which we compare.

In recent years, groundwater modeling is one of the most important topics in engineering and geoscience. Numerous mathematical models have been developed to solve groundwater problems [*Remson et al.*, 1971; *Wang and Anderson*, 1982; *Yeh*,

1986; *Hanna*, 1995]. Numerical models have often been used as effective tools to analyze groundwater systems. Generally speaking, physically based mathematical models are solved by finite-difference or finite-element methods and most of these models are distributed parameter models. The parameters used to characterize the groundwater numerical models are not directly measurable and have to be determined through parameter estimation processes. As a result, parameter calibrations in groundwater models are essential for successful modeling and the inaccuracy of parameter estimation may cause unreliable model output for future predictions or management purposes. The problem of parameter identification has been studied extensively during the past decades, and numerous approaches have been developed for solving this problem. Groundwater inverse methods have been reviewed by *Yeh* [1986], *Kuiper* [1986], *Ginn and Cushman* [1990], *Sun* [1994], *Sun et al.* [1995], *McLaughlin and Townley* [1996], *Hyun and Lee* [1998], *Carrera et al.* [2005], *Hill and Tiedeman* [2007] and *Hill et al.* [1998].

Various techniques have been developed to solve the parameter estimation problem. Manual calibration, which is also named trial-and-error method, has been popular and is a frequently used approach for model calibration. The hydrologists or the modelers specify the initial parameter values according to their experience and adjust the parameters based on comparing the model-simulated values with the observed values. The whole processes require a large amount of human time as well as

perception of the model. Thus, the manual calibration is very tedious and time-consuming.

To reduce computation time and human effort, automatic calibration that involves the use of an optimization method to search for the parameter set subject to a specified goodness-of-fit function were developed. Compared to manual calibration, automatic calibration requires much less human time and has higher possibility finding a better parameter set since more model simulations could be performed without worrying about saving human effort. To solve the automatic calibration problem, Gradient-based optimization methods, such as Gauss-Newton, gradient steepest descent, conjugate gradient, quasi-Newton, truncated-Newton, and Levenberg-Marquardt methods, have been widely used in groundwater model calibration. Previous researches have demonstrated the performance of applying gradient-based algorithms to groundwater calibration problems [*Lin and Yeh*, 1974; *Aral*, 1985; *Yeh*, 1986; *CHENG and Yeh*, 1992; *Olsthoorn*, 1995; *Heidari and Ranjithan*, 1998; *Hill et al.*, 1998; *Karahan and Ayvaz*, 2005]. The variations of the Gauss-Newton optimization approach were written into solution codes for applying into groundwater inverse problems, such as UCODE [*Poeter et al.*, 2005], iTough2 [*Finsterle*, 2007], PEST [*Doherty*, 2010]. The model independent Levenberg-Marquardt (LM) method based parameter estimation software PEST, which quantifies model-to-measurement misfit in the weighted least squares sense, has been widely used for environmental numerical model calibration. This software is efficient in terms of its model run requirements.

The major advantage of using these gradient-based algorithms is that these local optimization methods are computationally efficient at searching for local minima of the non-linear objective functions. However, groundwater inverse problems are typically highly nonlinear with multiple local minima. Thus, these gradient-based methods can be easily trapped into local optimum and cannot necessarily find the global optimum solution.

To solve this problem, researchers have developed a number of heuristic methods for automatic calibration. These classic methods include the Genetic Algorithm [*Goldberg*, 1989; *Lingireddy*, 1998; *Prasad and Rastogi*, 2001; *Giacobbo et al.*, 2002; *Tung et al.*, 2003], Artificial Neural Networks [*Karahan and Ayvaz*, 2006; 2008], tabu search [*Zheng and Wang*, 1996], simulated annealing [*Zheng and Wang*, 1996; *Tung et al.*, 2003], and ant colony optimization (ACO) [*Abbaspour et al.*, 2001]. In addition, there are some other heuristic methods were developed and implemented to inverse problems. For example, *Solomatine et al.* [1999] combine clustering and multi-start local search algorithms. *Duan et al.* [1993] proposed the shuffled complex evolution (SCE) algorithm and applied this global optimization method for calibrating watershed models [*Duan et al.*, 1994]. *Tolson and Shoemaker* [2007] proposed dynamically dimensioned search (DDS) algorithm and used this efficient global optimization method in watershed model calibration problems. *Hansen and Ostermeier* [2001] developed Covariance Matrix Adaptation Evolution Strategy (CMA-ES). This algorithm has been encapsulated into the PEST software package [*Doherty*, 2012]

with name CMAES_P. In this paper, our new optimization method is tested extensively against PEST derivative-based method and CMA-ES. The difficulty with heuristic methods is that they tend to require hundreds of thousands of function evaluations to obtain adequately good solutions, which is not very practical for some computationally expensive models. To improve the efficiency of heuristic methods, some hybrid global optimization methods are studied to enhance the performance of these heuristic algorithms. For example, *Agyei and Hatfield* [2006] introduces a hybrid global optimization algorithm which integrates the global search capacity of SCE with the high efficiency of gradient-based Levenberg-Marquardt (GBLM) method and tests this method with several inverse problems where parameters of a nonlinear numerical groundwater glow model are estimated. Although the hybrid methods have been recognized to be more powerful and robust than conventional nonlinear programming techniques, the problem of demanding many function evaluations still remains especially for computationally expensive groundwater models.

To overcome this problem, derivative-free optimization methods are introduced. These derivative-free methods utilize response surfaces in place of derivatives. Examples include UOBYQA [*Powell*, 2002], which has been applied to watershed model calibration by *Shoemaker et al.* [2007], NEWUOA [*Powell*, 2006; 2008], DFO [*Conn et al.*, 2006] and ORBIT [*Wild et al.*, 2009]. [*Mugunthan et al.*, 2005] apply a function approximation algorithm (FA-RS) [*Regis*, 2004] method to bioremediation models. *Shoemaker et al.* [2007] utilize a multi-start method in combination with a

derivative-free local optimization method ESRBF [*Regis and Shoemaker*, 2004] for watershed model calibration.

Because groundwater models vary by many orders of magnitude depending on what model, spatial discretization level and the aquifer size is chosen for the modeling case study, the required time for a single simulation could be different from a few minutes to several days. Therefore, investigating more effective and robust global optimization methods becomes the goal and is necessary if the groundwater numerical model is highly computationally expensive. An global optimization algorithm Stochastic RBF [*Regis and Shoemaker*, 2007] that we introduce in this paper uses a radial basis function (RBF) as a response surface instead of the quadratic surface used in earlier derivative free local optimization methods.

In this paper, we conduct the performance comparisons among Stochastic RBF, PEST derivative-based algorithm and CMAES_P with applications on groundwater numerical models based on real aquifer as well as test functions. The reason we select PEST to be the baseline method is that PEST has been widely used and recognized as a powerful tool for environmental model parameter estimation problems. We present the algorithm performance results in ways that are meaningful for modelers subject to a wide range of computational limitations.

The remainder of this study is organized as follows. Chapter 2 mainly introduces the problem of model calibration. The bench mark optimization algorithms utilized in

this study, and the Stochastic RBF algorithm is described in detail in Chapter 3. In Chapter 4, we introduce the aquifer to which the automatic calibration is applied and the results. The application to test functions and the results are described in Chapter 5. Conclusions and discussions are detailed in Chapter 6.

CHAPTER 2

THE PROBLEM OF MODEL CALIBRATION

2.1 Groundwater Flow Modeling

We consider calibration of a model of transient groundwater flow in an unconfined, heterogeneous, isotropic aquifer and we assume the groundwater is incompressible with constant density and viscosity. The equations used in this study to describe groundwater flow are

$$\frac{\partial}{\partial x}(hK\frac{\partial h}{\partial x})+\frac{\partial}{\partial y}(hK\frac{\partial h}{\partial y})\pm Q=\mu\frac{\partial h}{\partial t} \qquad x,y\in\Omega,\ t=0 \tag{1}$$

$$h(x,y,t)\big|_{t=0}=h(x,y,0) \qquad x,y\in\Omega,t=0 \tag{2}$$

$$h(x,y,t)\big|_{\Gamma_1}=h(x,y,t) \qquad x,y\in\Gamma_1,t\geq 0 \tag{3}$$

$$hK\frac{\partial h}{\partial \vec{n}}\big|_{\Gamma_2}=q(x,y,t) \qquad x,y\in\Gamma_2,t\geq 0 \tag{4}$$

where t is time (T); x, y are Cartesian coordinates (L); h is the hydraulic head (L), $h(x, y)$; K is the value of hydraulic conductivity (L/T); μ is specific storage coefficient of the aquifer (1/L); Q is the fluid sinks/sources term (1/T); Ω is the model domain; Γ_1 is Dirichlet boundary; Γ_2 is Neumann boundary; $\vec{n}$ denotes the normal to the boundary

Γ_2; *h(x, y, 0)* indicates the initial water table; *h(x, y, t)* is the water table on Dirichlet boundary at time *t*; and *q(x, y, t)* represents lateral flux of Neumann boundary.

Since the direction and rate of groundwater flow is determined by spatial or temporal variations in some hydrologic and hydro-geological parameters (e.g. *K* and μ), to apply the groundwater flow models, the knowledge of these hydrologic and hydro-geological parameters is required. Therefore, as one of the first steps in modeling study, field measurements of these parameters, such as pumping tests, are essentially point measurements providing an estimation of parameters for the area near the observation wells. However, the data from field measurements can only represent a small part of the study area in many cases because of the limited number of observation wells. Hence estimations of spatially distributed parameters with whole aquifer model rather than using point measurements become necessary. Hence the field measurements are used to establish the ranges of each parameter in each zone and the optimization-driven calibration is used to estimate the parameters in each zone within this range.

2.2 Formulation for Optimization

In practical groundwater model application, input parameters are never fully defined and are associated with a variety of uncertainties with limited knowledge of model inputs, even though many site measurements and studies already have been

made beforehand. Automatic calibration with optimization algorithms is a way to obtain the best values of model input parameters. In most cases, the objective function in the optimization is to minimize an error function which defines the discrepancy between model outputs and the observations.

Generally, calibration of parameters can be formulated as a box-constrained minimization problem as follows:

$$\min_{\mathbf{p}} f(\mathbf{p}) \tag{5}$$

subject to

$$p_{\min_k} \le p_k \le p_{\max_k}, k \in \{1, 2, \mathrm{K}\ , n\} \tag{6}$$

where, $\mathbf{p} \in \mathbf{R}^d$ denotes a vector of model parameters; $\mathbf{p_{max}}$ and $\mathbf{p_{min}}$ are vectors of the upper and lower bounds for parameters, which usually come from physically feasible range of parameters, prior information and expert experience; $h(\mathbf{p})$ is a groundwater simulation model as a function of parameter $\mathbf{p}$; $f(h(\mathbf{p}))$ is the error function (difference between observations and model outputs from $h(\mathbf{p})$) in calibration procedures as discussed in the following paragraphs. Determining the bounds of parameters can tremendously affect the efficiency of the optimization. If the ranges of parameters are narrowed down because of additional information of parameters, model parameter values may become less challenging to find. Thus incorporating as much

information and knowledge about the model input parameters as possible and constraining these ranges to be small is required before starting model simulation.

In an effort to compute the objective function $f(h(\mathbf{p}))$ for each set of potential parameter vector $\mathbf{p}$, the groundwater numerical model has to be run and the model output simulated. As discussed in the introduction, we mainly focus on the cases where the groundwater model is computationally expensive to simulate, which may take from many minutes to many hours to compute just one simulation in serial or parallel. As a result, only very limited number of function evaluations are allowed with a limited computational budget.

For the purposes of this study, simulation model output is denoted as $h_{i,j}^{sim}(\mathbf{K},\boldsymbol{\mu})$ that specifies the hydraulic head at well j and simulation time period i for given parameter sets $\mathbf{K}$, $\boldsymbol{\mu}$. The value $h_{i,j}^{obs}$ represents the corresponding observed hydraulic heads. The whole aquifer domain is divided into a set of N zones, thus each component of the model parameter (N-dimensioned vectors $\mathbf{K}$ and $\boldsymbol{\mu}$) is associated with one zone. The objective function seeks to find a good match between the observed and simulated hydraulic heads by minimizing their squared difference. There are a number of functions that can be used to specify the objective function such as total squared residual error, R^2, root mean squared error, maximum absolute error and Nash-Sutcliffe index (NSE) [*Nash and Sutcliffe*, 1970]. In this study, a nonlinear least squares function is implemented to evaluate the goodness-of-fit measures of

groundwater model calibration; the objective function for the optimization algorithm is to find the parameter sets that minimize the SSE.

The objective function formulated by total squared residual error (SSE) is defined as follows:

$$\text{Minimize} \quad SSE(\mathbf{K}, \boldsymbol{\mu}) = \sum_{j=1}^{T} \sum_{i=1}^{N} [\omega_{i,j} (h_{i,j}^{obs} - h_{i,j}^{sim}(\mathbf{K}, \boldsymbol{\mu}))^2] \tag{7}$$

where, T is the total simulation period of time; N is the total number of wells; SSE($\mathbf{K}$, $\boldsymbol{\mu}$) is the sum of squared error between observed and simulated hydraulic heads given the values of parameter sets ($\mathbf{K}$ and $\boldsymbol{\mu}$) to be calibrated in the groundwater numerical model. The lengths of these two vectors ($\mathbf{K}$ and $\boldsymbol{\mu}$) are both fourteen since each component of the parameter vectors is associated with one zone of the aquifer area and there are N=14 zones.

Groundwater models can be nonlinear, non-convex, non-smooth and even multimodal function of parameter values, so the corresponding inverse problems are very complicated to solve and the objective function has multiple local minima. The optimization processes requires a repeated simulation of a forward groundwater model in order to compute simulated hydraulic head using decision variables from parameter sets $\mathbf{K}$ and $\boldsymbol{\mu}$. Since the distributed parameter groundwater model problems are usually extremely computationally expensive which may require from couple minutes to several days for every single simulation, more effective optimization algorithms

producing better results for global optimization calibration problems with smaller number of simulations are needed.

CHAPTER 3

OPTIMIZATION ALGORITHM DESCRIPTION

The sections below discuss the three algorithms we used in all tests. Results for SCEUA_P algorithm that we didn't discussed in this paper are given in Figure 3b which shows the algorithm comparisons applying to Coarse Grid Case as well as the other three algorithms. Like CMAES_P, SCEUA_P is a global optimizer that encapsulated into PEST software suite. SCEUA_P implements the SCE (shuffled complex evolution) Algorithm developed by *Duan* [1991] and *Duan et al.* [1992; 1993; 1994]; UA means University of Arizona. SCEUA_P in Figure 3b has highest objective function value for same number of function evaluation, which is the worst in terms of algorithm efficiency among all four algorithms. As a result, we didn't test SCEUA_P for Fine Grid Case and hypothetical Case.

3.1 PEST Derivative-Based Algorithm

Model calibration was conducted using the Parameter Estimation (PEST). PEST [*Doherty*, 2010] is a model calibration program that is the most widely used and referenced automatic calibration method in groundwater analysis. The algorithm that is referred as PEST in this paper is the most widely used option in PEST that utilize the Gauss-Marquardt-Levenberg (GML) [*Marquardt*, 1963] algorithm which is a derivative-based algorithm. Many features are added to PEST, like the Automatic User Intervention (including truncated singular value decomposition method) that instructs

the algorithm to not perturb some most insensitive parameters for a certain number of iterations. PEST is very efficient to find the local optimal when the objective function is convex. PEST requires a significant amount of user input for the algorithm parameters such as maximum increment for derivative computation, step size, stopping criterion, etc.. However, it can be applied to many existing simulation models without accessing to models' source code, thus allowing simple calibration setup with an arbitrary model. Although the algorithm PEST will stop when a local optimum is found, the user can manually restart the algorithm at different starting points. In this paper, we implement PEST with many different starting points.

3.2 CMAES_P

CMAES_P is a "PEST compatible" [*Doherty*, 2012] implementation of CMA-ES (Covariance Matrix Adaptation Evolution Strategy), which is a global heuristic optimization algorithm that was developed by *Hansen and Ostermeier* [2001]. CMA-ES is an evolutionary algorithm for difficult non-linear non-convex optimization problems in continuous domain. It is typically applied to unconstrained or bounded constraint optimization problems, and search space dimensions between three and a hundred [*Hansen and Kern*, 2004]. The population size is λ and n is the dimension of parameter vector. In each iteration, λ random realizations of n-dimensional parameter vectors are generated, and the objective function is computed for each iteration. The n-dimensional covariance matrix used to produce parameter realizations alters

according to the objective function surface from sampling it during the optimization process in order to make the whole process more efficient. The manual for CMAES indicates that as the algorithm iterates, the neighborhood is narrowed such that parameter sets that are less likely to produce the global minimum are less likely to be considered. The narrowing is based on which solutions in previous iterations are selected as best. At the same time, the chances of being trapped in a local minimum of the objective function are also reduced randomly during the optimization process. After the λ model runs, μ lowest objective functions are selected to calculate a new parameter set. Weights can be added to the process of computing parameter set, and the weights determined by users according the situation. In this study, we use the "super linear" weights, which is an option described in the manual. Since CMA-ES is encapsulated in the package of PEST, the name of the algorithm has been changed into CMAES_P. Thus, we use CMAES_P in this paper to refer to the CMA-ES algorithm.

3.3 Stochastic RBF

Regis and Shoemaker [2007] introduced the Stochastic Response Surface (SRS) Method as a new framework for stochastic global optimization of expensive objective functions using response surface models. Two new global optimization methods were developed: Global Metric Stochastic Radial Basis Function (Global MSRBF) and Multistart Local Stochastic Radial Basis Function (Multistart Local MSRBF). In this paper, we used Multi start Local MSRBF. The key idea of these two global

optimization methods is to utilize radial basis functions (RBF) [*Powell*, 1992; 1999; *Buhmann*, 2003] as the response surface model to approximate the expensive objective function and thus reduce the number of function evaluations needed. Global MSRBF and Multistart Local MSRBF are both designed for continuous, multimodel and computationally expensive functions, especially if no derivative information is inexpensively available.

In Stochastic RBF, radial basis functions (RBF) are used to approximate the objective functions that are expensive to compute. The purpose of using RBF is to reduce the computational expense of an optimization problem by letting the RBF approximation identify candidate points that are unlikely to be highly fit before the actual simulations of $f(h(\mathbf{p}))$ are done.

We need to do some initial evaluations of the objective function $f(h(\mathbf{p}))$ at a series of points $\{p_i\}$ to get an initial response surface. There are several ways to choose the experimental design points **p** to construct the interpolant s_n. *Regis and Shoemaker* [2005] suggest that a Latin Hypercube Design (SLHD) can be employed, and we use this approach. For a d dimension problem, $2(d+1)$ symmetric SLHD points are used for initial surface.

Given n previously evaluated parameter set $p_1, p_2, \ldots, p_n$, a RBF interpolation model that approximates the objective function has the form

$$s_n(p) = \sum_{i=1}^{n} \lambda_i \phi(\|p - p_i\|) + t(p), \quad p \in R \tag{8}$$

where λ_i for $i = 1, \ldots, n$, $\|\bullet\|$ is the Euclidean norm. A cubic radial basis function with a linear tail $t(p)$ was chosen to fit the response surface for interpolation, thus giving $\phi(r) = r^3$.

CHAPTER 4

APPLICATION OF AUTOMATED CALIBRATION TO GROUNDWATER NUMERICAL MODEL

4.1 General Description of Study Aquifer

The Miyun-Huai-Shun watershed basin, as one of the major water supply resources of Beijing city, is located in the northern part of Beijing city, China (Figure 4.1). The mean annual precipitation and evaporation are 605 mm and 1120 mm, respectively. The precipitation varies seasonally and about 60 percent of the precipitations is concentrated in July and August which is typical for a temperate sub-continental monsoon climate regime.

The main aquifer area is 456 km^2 with three boundaries: eastern, western and south. According to previous studies of these three boundaries, they can be regarded as relatively impervious boundaries. Given that the horizontal dimension of this aquifer ranges from tens of kilometers while the depth varies only from tens to hundreds of meters, the groundwater problem in this aquifer has been simplified as a two-dimensional, unconfined, heterogeneous, isotropic, transient flow system. The model is governed by a partial differential equation as Eq.(1)., and the boundary conditions can be expressed as Eq.(2)-(4).

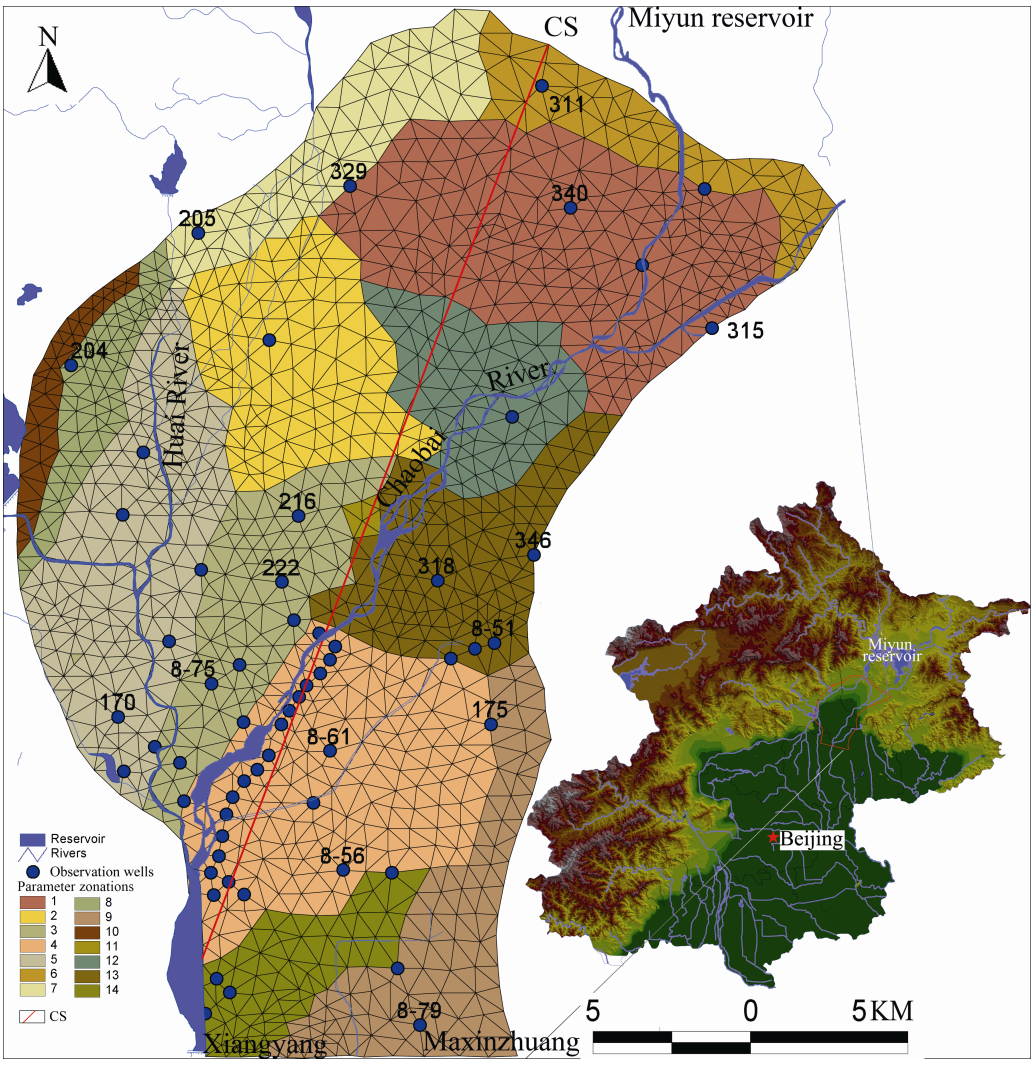

Figure 4.1: Miyun-Huai-Shun aquifer site and surroundings. Location with hydraulic head observation wells of the study area, zones of model for parameters [*Wei*, 2001]

Finite element method (FEM) was implemented to solve this groundwater problem. This FEM model, which is developed by *Wei* [2001] has been applied to several engineering projects with different regimes [*Wei*, 2001; *Wang et al.*, 2004].

4.2 Calibration and Study Cases

4.2.1 Parameterization

Parameterization is defined as a process of identifying the necessary parameters for a model. The groundwater model discussed in this paper is a spatially distributed numerical model. In the numerical calibration process, the parameters decided by parameterization will be systematically modified after a number of optimization iterations until good matches between the model simulated outputs and observations from known data are reached.

Usually, in a groundwater system, the zone structures are determined beforehand, and then the most important parameters required to be calibrated are selected. In our study, according to the lithological information and hydro-geologic characteristics from previous studies and research, the entire model aquifer can be divided into 14 zones in horizontal direction (Figure 4.1). The model parameters required to be determined by the optimization procedure vary according to different zones, but the parameters are assumed to be constant in the same zone.

In this study, two sets of model parameters, hydraulic conductivity and specific yield, were selected for the automatic calibration, since these two parameter sets are comparatively more important for groundwater model. Thus two parameters are taken into account in each zone, a total of (14×2=) 28 model parameters need to be

calibrated in the problem. The observed values of water levels were obtained from 46 observation wells (Figure 4.1) scattered in the entire aquifer. Therefore, the objective of automatic groundwater model calibration process in this study was to change the input parameter vectors (**K** and μ) until model outputs were close enough to the observed water levels of 46 observation wells.

4.2.2 Study Cases

Following our discussion in the aquifer description, we focus on the example cases that have been created in this study. The groundwater system discussed in this paper is modeled with Finite Element Method (FEM), and triangles grid were used to discretize the aquifer domain. According to the theory of FEM, up to some limit the finer the grid is, the more accurate FEM performs. To show the performance of two optimization algorithms in different grid size FEM models, two different scaled meshes were applied for the entire model area. The "Coarse Grid" has 478 nodes and 871 triangular elements. The individual grid area ranges from 0.059 km^2 to 1.71 km^2. The "Fine Grid" has 1239 nodes and 2337 elements (Figure 4.1) with a range of individual grid are from 0.032 km^2 to 0.5 km^2. "The Coarse Grid" is essential a "surrogate model" of original model, and it reduces lots of computation cost per simulation. 46 observation wells are irregularly scattered in the aquifer domain (Figure 4.1), which are used to observe ground water levels near wells.

In order to demonstrate the effectiveness of our algorithm, we selected three cases and compared the performance of these algorithms. In Case 1, the "Fine Grid" groundwater model was used and all 28 parameters (i.e., **K** and μ) were required to calibrate. In Case 2, we applied the "Coarse Grid" groundwater model with all 28 parameters calibration as well. These two cases were both based on the real aquifer using the observed values of water levels from 46 observation wells, whereas Case 3 was a hypothetical example which used "Fine Grid" model with synthetic observation data for calibration by simulating the "Fine Grid" model. Case 3 was created to consider a case with greater variability in hydraulic conductivity and specific yield among the zones in this aquifer. We will discuss later more details about the synthetic case.

In order to make the optimization process more efficient, more information and knowledge should be incorporated into the bounds of hydro-geologic parameters. For the first two cases, since they are based on real data, the bounds for model parameters (**K** and μ) should be based on physically feasible range of each parameters, prior information and expert experience. The available pumping tests demonstrate that the ranges of hydraulic conductivity and specific yield are from 120 m/d to 270 m/d, 0.12 to 0.24, respectively. However, we used slightly broader ranges than pumping tests since the pumping tests were taken in limited areas and the ranges from those tests might not be exact enough. The aquifer parameter bounds are shown in Table 4.1.

Table 4.1: Range of Calibrated Parameters for the Miyun-Huai-Shun Groundwater Aquifer

Zone	K_i (m/d)		μ_i	
i	Lower Bound	Upper Bound	Lower Bound	Upper Bound
1	50	800	0.10	0.35
2	50	800	0.01	0.30
3	50	800	0.10	0.30
4	20	800	0.10	0.30
5	50	500	0.10	0.30
6	50	800	0.10	0.30
7	50	800	0.10	0.30
8	30	500	0.10	0.30
9	30	800	0.10	0.35
10	50	500	0.10	0.35
11	20	500	0.01	0.35
12	50	800	0.01	0.35
13	50	800	0.10	0.35
14	20	500	0.10	0.35

4.2.3 Hypothetical Illustrative Example

In this study, we created a study case where the algorithms would be tested on a related but probably more difficult calibration problem because the true values of hydraulic conductivity and specific yield are more variable than ones in the Miyun-Huai-Shun aquifer.

The case is adapted from "Fine Grid" Beijing groundwater numerical model [*Wei*, 2001] and involves flow along the same domain shown in Figure 4.1. In a real Case, observed data are often obtained by pilot tests or studies from previous records. Thus, data such as hydraulic heads is usually available for calibration of unknown parameters. Since a hypothetical Case is considered here, the measured data for model calibration were obtained by simulating the groundwater model. The same as the "Fine Grid" model, the hypothetical aquifer is divided into 14 zones of constant hydraulic conductivity and specific yield. Therefore, 28 parameters in total were considered unknown and needed to be calibrated. The bounds of parameters for the hypothetical case are based on the feasible ranges of hydraulic conductivity and specific yield, but they are much broader than the real Cases since we assume that no information about the parameters bounds are known. Thus, there might be more local minima in the hypothetical case because of the large ranges of parameters. The lower and upper bounds for K and μ are 5 m/d~8000 m/d and 0.01~5, respectively. According to the previous research, the initial values for the parameters of hydraulic conductivity and specific yield were obtained from the analyzed results of well test data [*Wang et al.*, 2004]. The values of parameters are given in Table 4.2. In order to obtain the appropriate predetermined parameter values, we plotted the parameter values of Table 4.2 to try to find the relationship between hydraulic conductivity and specific yield. Figure 4.2 illustrates that there is a strong positive linear relationship between these two parameter sets; thus, reasonable specific yield can be obtained according to linear regression analysis once we randomly assign values to hydraulic conductivities of 14

zones. The "True" model parameter values are shown in Table 4.3. The measured hydraulic heads were generated by simulating the "Fine Grid" groundwater numerical model using the predetermined values of those 28 parameters. These predetermined values will be considered as the true values for these parameters. Since the measured data should be noisy because of measurement error, the noise of data set available for calibration needed to be considered. In order to simulate errors, normally distributed random noise with a standard deviation of 25% of the true values was incorporated to the simulated hydraulic heads that were used for calibration. Thus, these hydraulic heads were used as synthetic observations in assessing performance of algorithms in automatic calibration.

Table 4.2: Values for Parameters of Hydraulic Conductivity and Specific Yield from Previous Research of Analyzed results of Well Test [*Wang et al.*, 2004]

Zone i	K_i (m/d)	μ_i
1	225	0.17
2	200	0.15
3	175	0.14
4	188	0.14
5	130	0.12
6	130	0.09
7	100	0.09
8	85	0.09
9	60	0.06
10	50	0.03
11	8	0.001
12	215	0.165
13	195	0.15
14	110	0.11

Table 4.3: True Parameter Values of Hypothetical Case (Case 3)

Zone i	$\boldsymbol{K}_i$ (m/d)	$\boldsymbol{\mu}_i$
1	3234	2.26
2	776	0.57
3	1060	0.76
4	7537	5.29
5	12	0.02
6	4604	3.23
7	8	0.01
8	10	0.02
9	2829	1.99
10	12	0.03
11	13	0.03
12	349	0.26
13	1356	0.96
14	5195	3.66

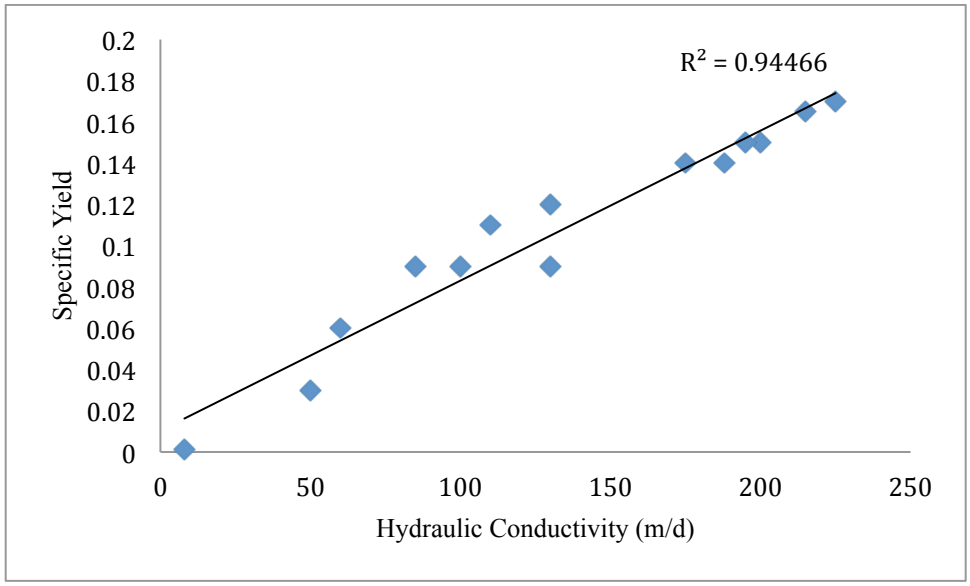

Figure 4.2: Relationship between Hydraulic Conductivity and Specific Yield according to Previous Research from [*Wang et al.*, 2004] when used in Construction of Case 3

4.3 Numerical Results

In the first two Cases, to calibrate model parameters, 14 years (1985 to 1999) of observation data at 46 wells throughout the model domain were used as the references for fit with the calculated results. For the Coarse Grid groundwater model (Case 2), each simulation took about 8.5 seconds on an Inter® Core® 2 Quad CPU Q6600@ 2.40GHz computer, and 4 minutes and 18 seconds for the Fine Grid model (Case 1) and Hypothetical model (Case 3).

Each algorithm was run 20 times for up to 500 function evaluations for every Case. The reason for doing 20 trials for each algorithm is that the optimization algorithms we used are all stochastic. As a result, different users of an algorithm could obtain a different sequence of solutions for a given run of the optimization algorithm.

For statistical comparison of optimization algorithms, doing multiple trials for each algorithm is necessary and is also good for characterizing the variability in performance. Randomly generated starting points that were different from trials were used for PEST and CMAES_P. Regis and Shoemaker [2007] suggested the use of a Latin Hypercube Experimental design (LHD) for fitting the initial response surface of Stochastic RBF. Thus, we used a different set of independently generated Latin Hypercube points in each trial for Stochastic RBF.

To compare the optimization algorithms, we examined the following characteristics of these algorithms: (1) efficiency in giving good objective function for a given number of function evaluations; (2) performance for different statistical criteria; (3) variability of solutions in multiple trials; (4) improvements if more function evaluations are performed. We developed a new hybrid algorithm that combines Stochastic RBF and PEST for model parameter calibration.

4.3.1 Analysis of Average Results for Multiple Cases

The performance of the algorithms is compared in Figure 4.3 by using plots of the average best objective function value versus the number of function evaluations. Average best value indicates the average over 20 trials of the best objective function solution obtained in the given number of evaluations. The lowest curves are the best since the goal of optimization is to reduce the differences between the observed data and simulated data, and hence have the lowest objective value. Thus, the smaller the average objective values are, the better optimization results we get.

Stochastic RBF is clearly superior to all the algorithms as it has the fastest drop and lowest average SSE at the same number of function evaluation in comparison to all other algorithms. CMAES_P algorithm performs well for all three cases, being the second most efficient algorithm. PEST is inferior in performance to the Stochastic RBF and CMAES_P as evidenced by slower drop and higher average objective function values with increase in function evaluations.

Stochastic RBF in hypothetical case has an even better performance than other two Cases since hypothetical case is designed to be more global. Hydraulic conductivity and specific yield vary more in Case 3 so this possibly results in more variability in the objective function surface. We note that for the most difficult problem the PEST algorithm is not able to reduce the objective function hardly at all

after the first 10 steps and the best value it finds (3752.15) is about third times as large as the value found by Stochastic RBF (1125.74).

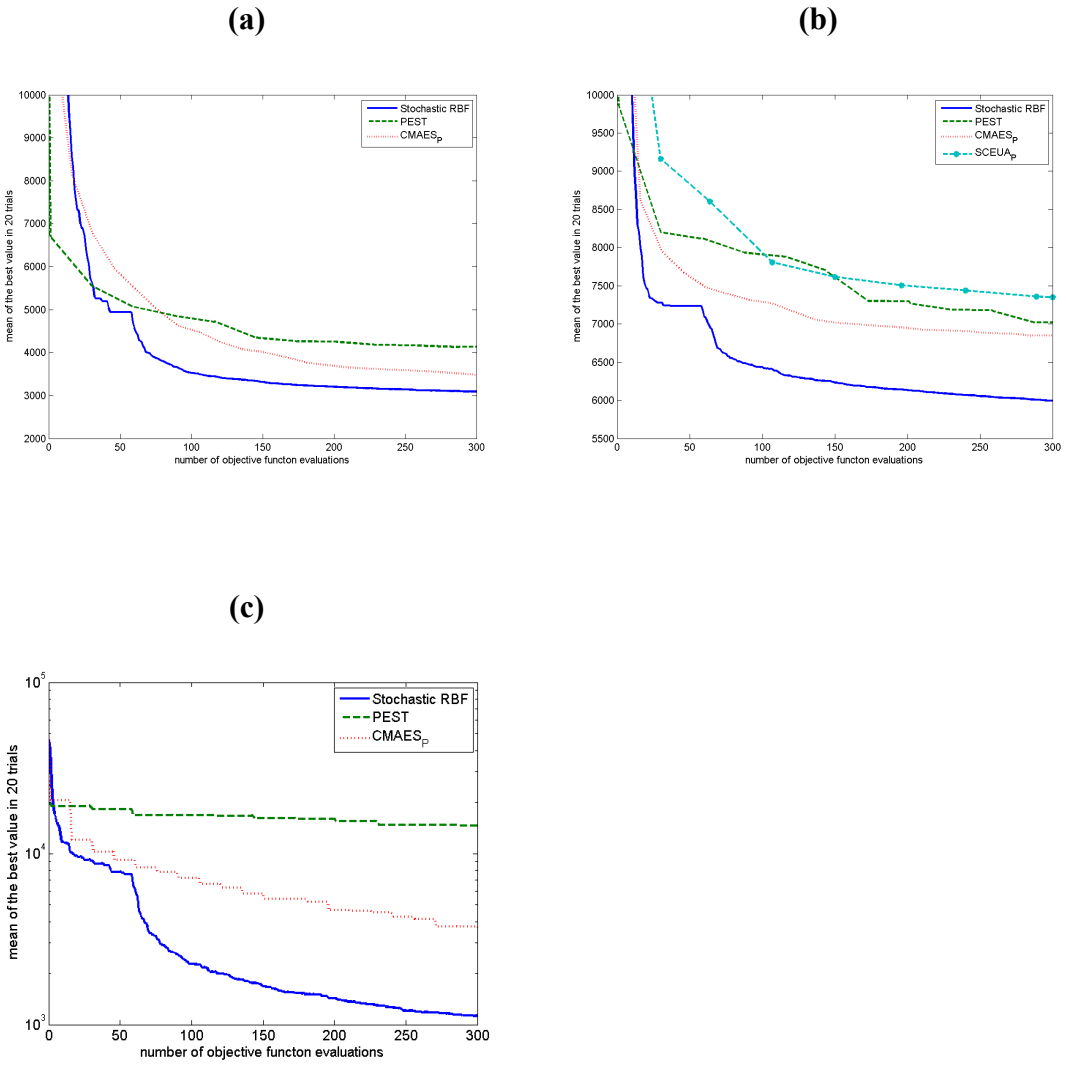

Figure 4.3: Best objective function value for each algorithm for 300 evaluations averaged over 20 trials: (a) Case 1: Fine Grid model for MHSA. (b) Case 2: Coarse Grid model for MHSA. (c) Case 3: Hypothetical Case. The lowest curve is the best. (MHSA is the model for the Miyun-Huai-Shun Aquifer)

4.3.2 Model Simulation Results

Figure 4.4 plots the simulated hydraulic head compared to the observed head during the calibration period as well as validation period by Stochastic RBF, PEST and CMAES_P of the fine grid groundwater model. The validation period began in January, 2000 and ended in December, 2002 with a monthly time step are used to validate the calibrated model. The statistics of validation results are shown later in Section 4.3.6. The simulations represent the effectiveness of the methods by looking at the parameter solutions found in middle ranking of the twenty trials. Hence, the simulations are based on the 10^{th} best parameter sets for all three algorithms. The 45 degree line is a measure of a good agreement between the simulated and measured hydraulic heads. If the simulated date is the same as the observed data, then the data point should be right on this line. If the residuals are mostly below the 45 degree line, it means the algorithm underestimates the hydraulic heads. Otherwise, the algorithm overestimates the hydraulic heads. Figure 4.4 indicates a worst agreement between the simulated and measured hydraulic head data of PEST evidenced by some biased data points in the upper right corner above the 45 degree line. Although the difference between the performance of Stochastic RBF and CMAES_P is not tremendous, we can still observe that the data points on the upper right corner of Figure 4.4(c) are still slightly above the 45 degree line.

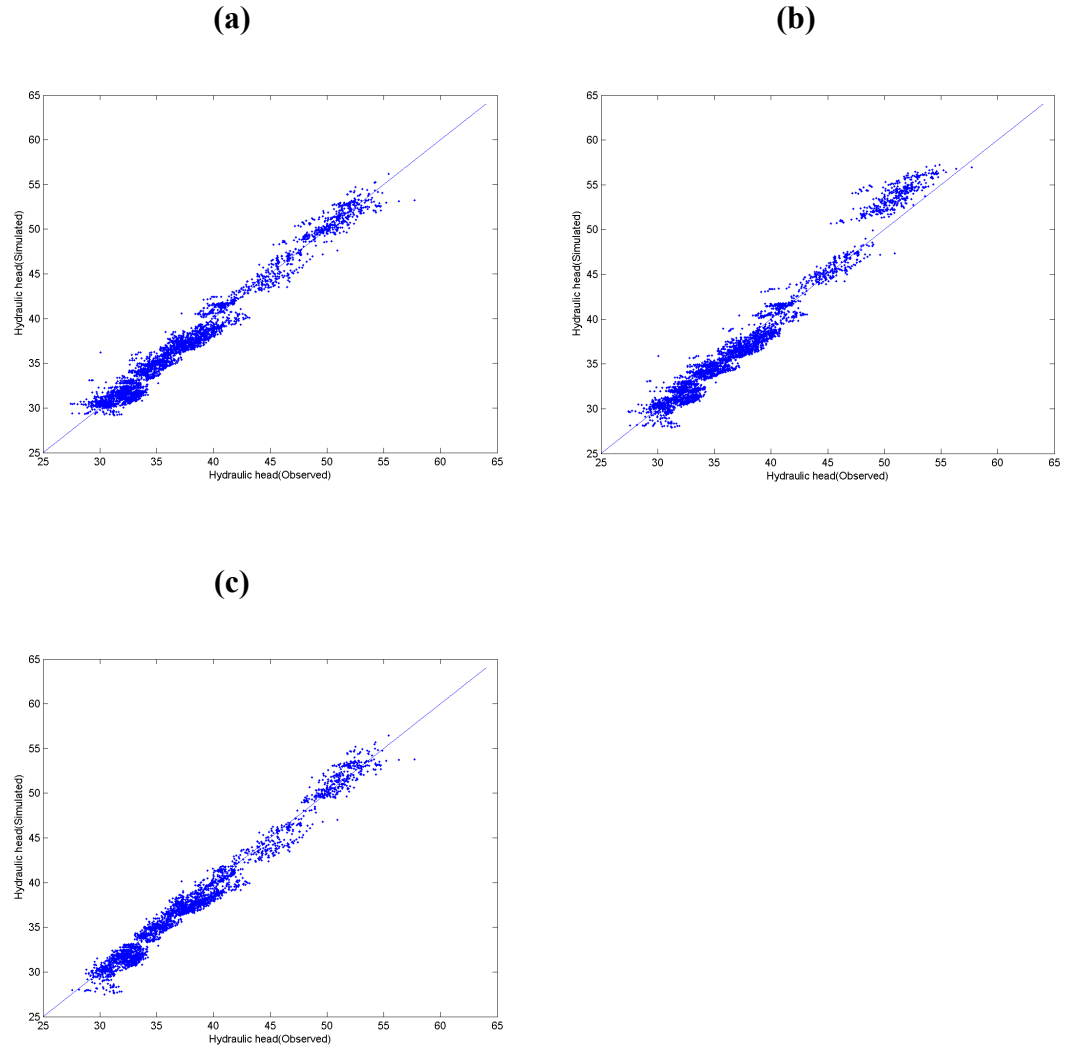

Figure 4.4: Simulated hydraulic head vs observed hydraulic head from fine grid Beijing Groundwater model with parameter calibrated by (a) Stochastic RBF, (b) PEST and (c) CMAES_P respectively, after 300 simulations (calibration period + validation period).

4.3.3 Variability in Multiple Trials

Although Figure 4.3 shows the average values of objective function, the variability in the solution produced by each algorithm cannot be demonstrated only by the mean objective function values. In a real calibration practice, probably only one algorithm will be applied to calibrate because the objective functions for environmental models are usually extremely expensive to calculate. Therefore, it is particularly important to choose an algorithm that can produce better results consistently, which implies the mean is low and the variance is small. A smaller variability with a low mean of objective function value suggests a more reliable algorithm for producing a good calibration in any given trial. Thus, an algorithm that produces a good solution consistently is clearly a superior choice to an algorithm that has a sizeable chance of producing a poor solution. The mean and standard deviation of objective function value for the best solution at the end of 300 function evaluations for all three Cases and algorithms are shown in Table 4.4.

In order to demonstrate the variability of these algorithms, the best solution at the end of 300 function evaluations of each algorithm are plotted into box plots. Each algorithm and Case was run for 20 trials. A total of 300 objective function evaluations were considered in each trial for all algorithms.

The Stochastic RBF has the best median and smaller variability in the objective function values obtained in 20 optimization trials for all these cases as indicated by the

graphical box plot comparisons shown in Figure 4.5. Figure 4.5a, 4.5b, and 4.5c show box plots for Case 1, Case 2, and Case 3, respectively. The box plot shows the median, interquartile range, and outliers based on the 20 trials for each algorithm.

As can be seen in Figure 4.5, stochastic RBF has the smallest spread (more reliable) when compared to the other two algorithms for the SSE objective function formulation. PEST clearly is the worst among all three algorithms as indicated by their large spreads, with CMAES_P performing slightly better than PEST.

In addition, both PEST and CMAES_P have outliers while there is no outlier in Stochastic RBF. The outlier value of PEST is substantially large in all the Cases because PEST is not a global optimization algorithm and the performance of PEST highly depends on the initial starting points. If the initial solutions assigned to PEST are close to local minima, PEST can be easily trapped into local minima without moving to the global optimal solution. Especially in the hypothetical case (Figure 4.5c), which was designed to be the most difficult problem among the three cases, PEST has an extremely large spread than CMAES_P and Stochastic RBF. Even though CMAES_P has a much smaller spread than PEST, it is still greater than the spread of Stochastic RBF. Further, Stochastic RBF, as can be seen in Figure 4.5c, has the lowest mean for all the three Cases. The same as the variability comparison, PEST, which has the largest mean, performs the worst.

Table 4.4: Average and Standard Deviation of the Objective Function Value Produced by Each Algorithm and Each Case after 300 Function Evaluations based on 20 trials

Algorithm	Case 1 (Fine Grid)		Case 2 (Coarse Grid)		Case 3 (Hypothetical)	
	Average	SD	Average	SD	Average	SD
Stochastic RBF	3091.69	195.4	5994.31	103.9	1125.74	396
PEST	4132	1364	7021.77	1372.8	14519.12	7478.02
CMAES_P	3519.21	339.3	6846.4	743.9	3752.15	4591.7

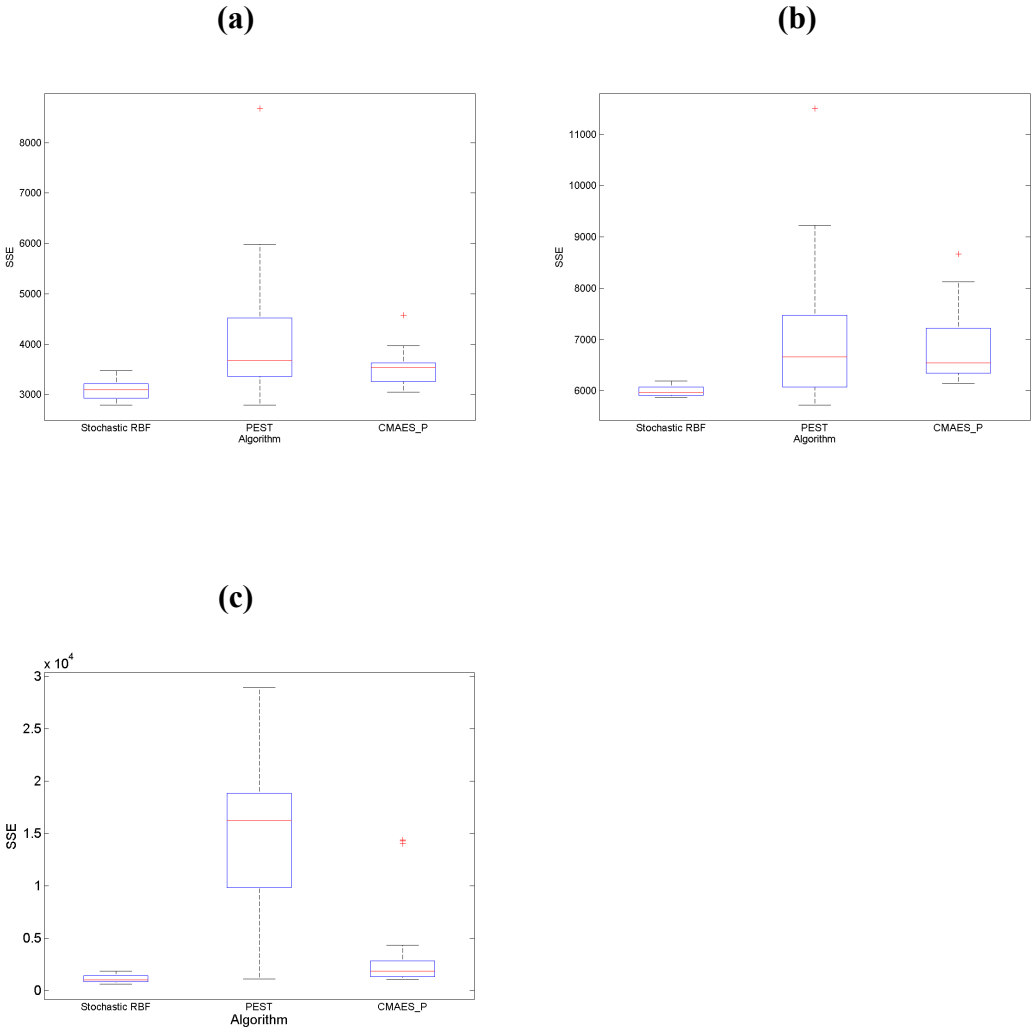

Figure 4.5: Box plot of best solution for each algorithm based on 20 trials after 300 function evaluations: (a) Case 1: Fine Grid model for MHSA. (b) Case 2: Coarse Grid model for MHSA. (c) Case 3: Hypothetical Case. Smaller box means more consistency across all trials. (MHSA is the model for the Miyun-Huai-Shun Aquifer)

4.3.4 Results for More Iterations

We showed the algorithm performance over 3 Cases for 300 function evaluations in section 4.3.1, and the results indicate that Stochastic RBF is superior to PEST and CMAES_P up to 300 function evaluations. In this section, Figure 4.6 illustrates the impact on average results (over 20 trials) of increasing the number of function evaluations to 500. As can be seen in Figure 4.6, the results are very consistent with the results in Figure 4.3, and the results also illustrates that most of the convergence was achieved before 300 iterations, with Stochastic RBF dropping more quickly than the other methods. After 300 function evaluations, all three algorithms in fine grid case and hypothetical case didn't improve much as can be illustrated by Figure 4.6a and 4.6c. PEST improved a lot between 300 to 500 iterations in the coarse gird case because of the coarse model is not as global as the other two Cases, it's still not as good as Stochastic RBF at the end of 500 iteration. Table 4.5 lists the averages and standard deviations of the objective function value of all three Cases and all algorithms after 500 function evaluations.

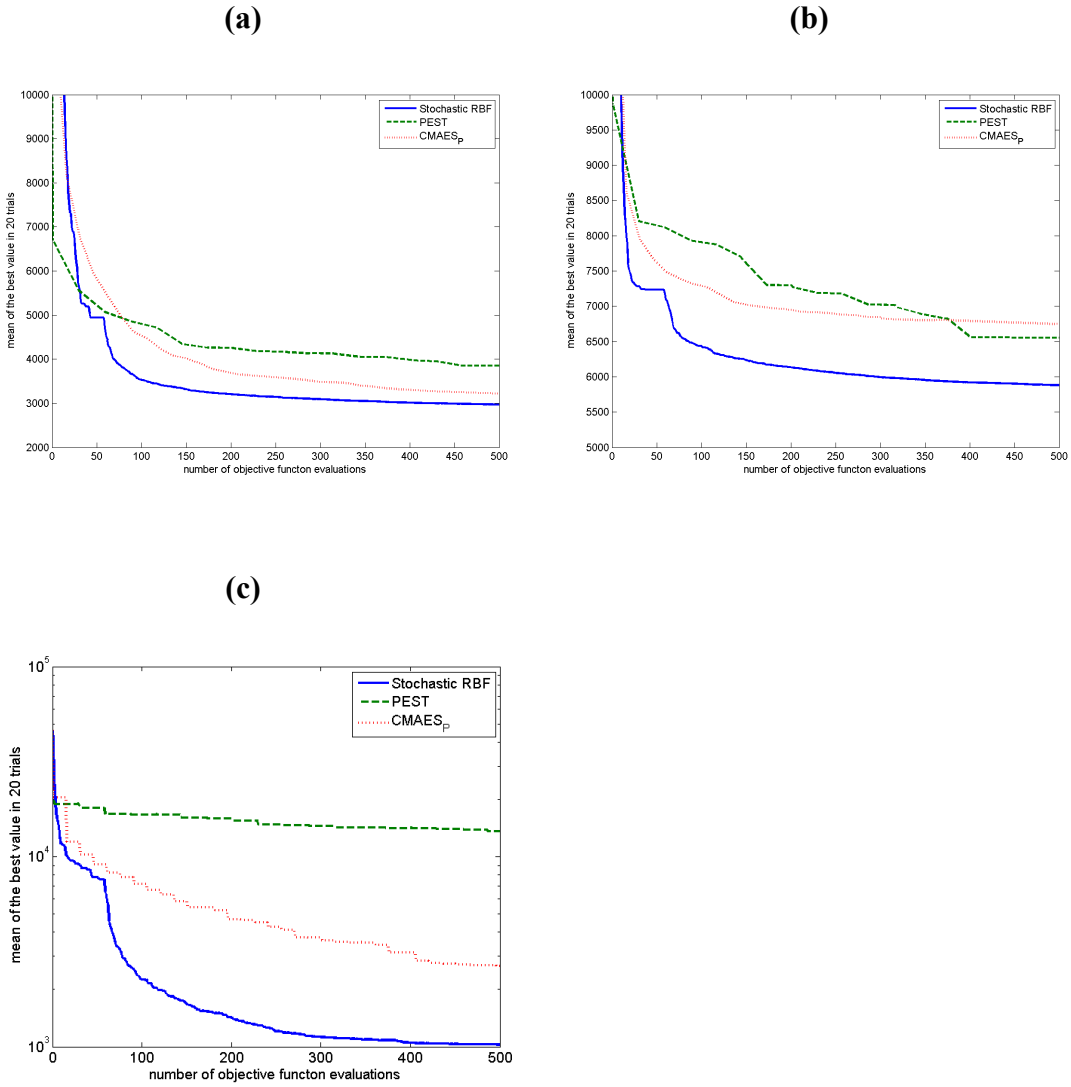

Figure 4.6: Average best objective function value for each algorithm formed as function of number of function evaluations of $f(h(\mathrm{x}))$ over 20 trials up to 500 simulations: (a) Fine Grid Case. (b) Coarse Grid Case. (c) Hypothetical Case.

Table 4.5: Average and Standard Deviation of the Objective Function Value Produced by Each Algorithm and Each Case after 500 Function Evaluations based on 20 trials

Algorithm	Case 1 (Fine Grid)		Case 2 (Coarse Grid)		Case 3 (Hypothetical)	
	Average	SD	Average	SD	Average	SD
Stochastic RBF	2968.10	151.9	5879.70	67.1	1026.08	332.2
PEST	3850.50	1376.5	6551.67	977.7	13631.86	8366.8
CMAES_P	3221.93	187.7	6747.09	801.3	2657.93	3888.5

4.3.5 Combination of Stochastic RBF and PEST

In this paper, we proposed a new Hybrid method by combining of Stochastic RBF and PEST Derivative-Based Algorithm, and compared the results of using this method to the results by Stochastic RBF. In order to combine Stochastic RBF and PEST, the best parameters from Stochastic RBF after 100 iterations were selected as the starting points of PEST Derivative-Based Algorithm. With the initial parameter sets from Stochastic RBF, PEST was run for 200 function evaluations. In this paper, since each of the algorithms were tested based on 20 trials with different starting parameter sets in each of the Cases, 20 trials of PEST were run with the corresponding initial parameter sets from 20 trials of Stochastic RBF.

The reason of combining these two algorithms is because of the characteristics of these two algorithms. Stochastic RBF is an exceptionally efficient global optimization

and PEST Derivative-Based Algorithm is prominent for its high-speed in finding the local minimum in computation expensive models especially when the number of parameters that need to be calibrated is large. As can be seen in Figure 4.3 and Figure 4.6, Stochastic RBF has the fastest drop than the other algorithms before 100 iteration. As a result, starting with applying Stochastic RBF first will expedite PEST finding the global minimum of the objective function given that PEST is fast in solving a local optimization problem. However, the performance of the combining method (combination of Stochastic RBF and PEST) depends on the optimization problem. If the problem is too global for Stochastic RBF to obtain a comparatively good solution after 100 iterations, the combining method could fail to find a better solution than Stochastic RBF since PEST doesn't perform as well as Stochastic RBF in a highly global optimization problem, which can be illustrated in Figure 4.7.

Figure 4.7 lists the performance of Stochastic RBF and the combination method in fine grid, course grid and the hypothetical grid Cases. In the coarse grid Case (Figure 4.7b), combination method of Stochastic RBF and CMAES_P is also compared to the other two methods, and the result shows that the combination method of Stochastic RBF and PEST performs the best among the three methods with the lowest average objective function value at the end of 300 iterations, whereas combining Stochastic RBF and CMAES_P performs the worst. The combining method doesn't work any better than Stochastic RBF in the fine grid Case (Figure

4.7a) and is has much higher objective function value in the hypothetical Case (Figure 4.7c) which was designed as the most global optimization problem.

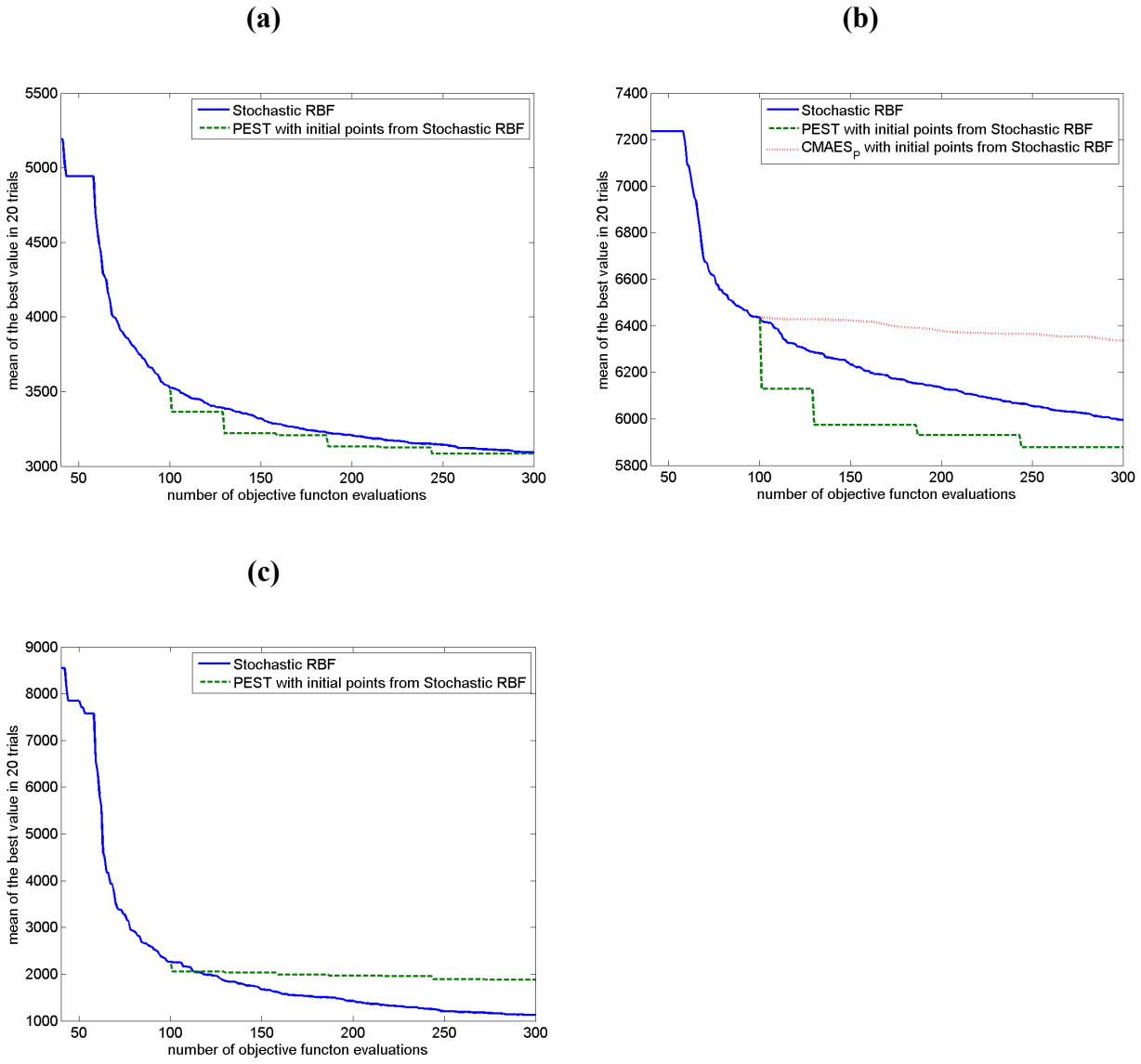

Figure 4.7: Comparison of algorithm performance including method of combining Stochastic RBF and PEST (start PEST with average parameters from 100th iteration of Stochastic RBF over 20 trials): (a) Case 1: Fine Grid model for MHSA. (b) Case 2: Coarse Grid model for MHSA. (c) Case 3: Hypothetical Case. The lowest curve is the best. (MHSA is the model for the Miyun-Huai-Shun Aquifer)

4.3.6 Multiple Performance Criteria

The goodness-of-fit describes how well the model outputs fit the observations. Measures of goodness-of-fit summarize the discrepancy between the values generated by model and the observed values. In order to evaluate the goodness-of-fit of the model simulation results, we applied three statistical indices: the Sum Square Error (SSE), the coefficient of determination (R^2), the Mean Absolute Error (MAE) and the Root Mean Square Deviation (RMSD). The equation of SSE is given in Equation (7), and the other three criteria will be explained in this section.

R^2 is a statistic that will give some information about the goodness-of-fit of a model. In theory, a R^2 of 1.0 indicates that the regression line perfectly fits the data. R^2 is defined as:

$$R^2 = 1 - \frac{SSE}{TSS} = 1 - \frac{\sum_{i=1}^{N} (h_i^{obs} - h_i^{sim})^2}{\sum_{i=1}^{N} (h_i^{obs} - \overline{h_i^{sim}})^2} \tag{8}$$

Where, TSS is the total sum of squares; N is the total number of measured data; h_i^{obs} and h_i^{sim} are observed hydraulic head data and simulated hydraulic head data, respectively; $\overline{h_i^{obs}}$ is the average of measured hydraulic head data.

In statistics, the MAE is a quality used to measure how close the simulated or predicted data are to the real data. Since it takes the absolute value of the error, the

lower the MAE is, the better match we get. However, the best MAE result does not necessarily mean the corresponding SSE result is the best. The MAE is given by

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|(h_i^{obs} - h_i^{sim})\right| \tag{9}$$

The RMSD is the square root of the averaged squared differences between observed and simulated hydraulic heads. The formula is

$$RMSD = \sqrt{MSE} = \sqrt{\frac{SSE}{N}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(h_i^{obs} - h_i^{sim})^2} \tag{10}$$

Table 4.6 in this section summarizes the results of the above four measures of goodness-of-fit for both the original calibration and validation period of the three Cases. In Table 4.6, 20 trials are performed for each combination of algorithm type and Case and the reported values for each of the four criteria are averaged over the 20 trials. For each Case, all the algorithms are evaluated based on their performance in 300 function evaluations.

The validation period began in January, 2000 and ended in December, 2002 with a monthly time step used to validate the calibrated model. In fact, the validation is a more stringent evaluation of a model since the parameters are not allowed to be adjusted to fit the real data. The statistics of validation can be used as a measure of how reliable the calibrated model could be for the future forecasts. In other words, a calibrated model which has good validation performance could be more likely to

produce reliable results on future conditions. All the validation statistics are also shown in Table 4.6.

Table 4.6 illustrates that Stochastic RBF finds the best solution for Case 1 and 3 on the calibration data set while PEST is significantly worse than any other algorithms in all three Cases. In Case 2, the mixed method is slightly better than Stochastic RBF and the possible reason we have stated in section 4.3.5. As indicated by the equations (8)-(10), the set of parameter values that optimizes the SEE also optimizes R^2 and RMES. However the percentage change in each of these quantities associated with a parameter changes will be different for each of these goodness-of-fit measures. Since R^2 includes a term that divides SSE by the variance, the differences in R^2 values are relatively small between algorithms in each Case while the percentage differences in SSE are quite large between the different algorithms.

Table 4.6: Multiple Performance Criteria results for all three Cases as well as all algorithms

Cases	Optimization Method	Calibration				Validation			
		SSE	R^2	MAE	RMSE	SSE	R^2	MAE	RMSE
Case 1 Fine Grid	Stochastic RBF	3091.7	0.9746	1.0024	0.7869	458.5	0.9753	0.9107	0.7231
	CMAES_P	3480.5	0.9714	1.0629	0.8376	532.2	0.9713	0.9806	0.7823
	PEST	3465.5	0.9663	1.1617	0.9073	666.5	0.9641	1.0758	0.8574
	Stochastic RBF + PEST	3094.9	0.9745	1.0026	0.7849	464.8	0.9750	0.9161	0.7310
Case 2 Coarse Grid	Stochastic RBF	5994.3	0.9632	1.3966	0.9534	1652.4	0.9285	1.7296	1.2019
	CMAES_P	6846.4	0.9607	1.4421	0.9929	1752.3	0.9241	1.7808	1.2462
	PEST	7021.0	0.9569	1.5056	1.0274	1963.8	0.9150	1.8716	1.3212
	Stochastic RBF + PEST	5877.1	0.9639	1.3829	0.9385	1587.4	0.9313	1.6954	1.1747
Case 3 Hypothetical Example	Stochastic RBF	1125.7	0.9809	0.5967	0.4088	185.1	0.9794	0.5725	0.3822
	CMAES_P	3752.15	0.9381	0.9598	0.6161	678.5	0.9246	0.9474	0.5877
	PEST	14519.1	0.7532	2.0703	1.2353	2740.7	0.6956	2.1252	1.2206
	Stochastic RBF + PEST	1729.2	0.9706	0.7012	0.5001	249.2	0.9723	0.6407	0.4540

CHAPTER 5

APPLICATION OF AUTOMATED CALIBRATION TO TEST FUNCTIONS

5.1 Test Function Problems Setup

We compared the performance of the Stochastic RBF, CMAES_P and PEST derivative based algorithm on two synthetic test function calibration problems. The test functions that we used in this paper are Ackley function and Hartman6 functions. Both of the test functions are unconstrained global problems designed for examining the performance of global optimization methods. In order to adapt these test functions to be suitable for calibration, the inverse problems of corresponding test functions were established. Certain known parameters from test functions that were assumed to be the decision variables that need to be calibrated were selected. Similar as constructing the hypothetical Case in section 4.2.3, "observed data" were obtained by simulating test functions using the real parameters. After adding noise to the simulation output, observed data was created and used as synthetic observations in assessing performance of algorithms in automatic calibration.

The Ackley function [*Ackley*, 1987] is a two-dimensional global test function. The function is

$$f(x_1, x_2) = -a\exp(-b\sqrt{\frac{(x_1^2 + x_2^2)}{2}}) - \exp(\frac{\cos(cx_1) + \cos(cx_2)}{2}) + a + \exp(1) \qquad (11)$$

where, *a*, *b* and *c* are equal to 20, 0.2 and 2π, respectively, and they were selected to be parameters to be calibrated. 100 observation data were generated through simulating the test function once. The lower bounds of parameters a, b and c are [0, -5, 10] and the higher bounds are [100, 5, 30].

The Hartman6 function [*Dixon and Szego*, 1978] is a six-dimensional function with multiple local minima. The function is

$$H(x) = -\sum_{i=1}^{4} \mathbf{a} \exp\left[-\sum_{j=1}^{6} B_{i,j}(x_j - Q_{i,j})^2 \right] \tag{12}$$

where

$$\mathbf{a} = [1, 1.2, 3, 3.2]$$

$$\mathbf{B} = \begin{bmatrix} 10 & 3 & 17 & 3.5 & 1.7 & 8 \\ 0.05 & 10 & 17 & 0.1 & 8 & 14 \\ 3 & 3.5 & 1.7 & 10 & 17 & 8 \\ 17 & 8 & 0.05 & 10 & 0.1 & 14 \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} 0.1312 & 0.1696 & 0.5569 & 0.0124 & 0.8283 & 0.5886 \\ 0.2329 & 0.4135 & 0.8307 & 0.3736 & 0.1004 & 0.9991 \\ 0.2348 & 0.1451 & 0.3522 & 0.2883 & 0.3047 & 0.6650 \\ 0.4047 & 0.8828 & 0.8732 & 0.5743 & 0.1091 & 0.0381 \end{bmatrix}$$

The elements in matrix B were selected to be the parameters for inverse calibration problem of Hartmann6 test function, thus the number of parameters is 24. The bounds are from -100 up to 300 for each of the elements in Matrix B.

5.2 Results

Twenty independent trials with a maximum of 500 function evaluations were performed using Stochastic RBF, CMAES_P and PEST. The progress of the algorithms toward an optimal solution is compared in Figure 5.1 by using a plot of the average best objective function value versus the number of function evaluations. As is shown in Figure 5.1, Stochastic RBF is converging to superior solutions at a much quicker rate than PEST and CMAES_P. The performances of CMEAS_P in both cases are better than PEST. CMAES_P is close to Stochastic RBF in Figure 5.1a, whereas in Figure 5.1b, the performance of CMEAS_P is much worse than Stochastic RBF. The reason might be that the dimension of the Hartmann6 function case is 24, which is 8 times larger than Ackley function case. As dimension increases, calibration problems are getting gradually difficult to solve and become more global. As a local optimization algorithm, PEST will be easily trapped in local minima as problem getting more global, resulting in the unsatisfactory performance in both two cases.

(a) **(b)**

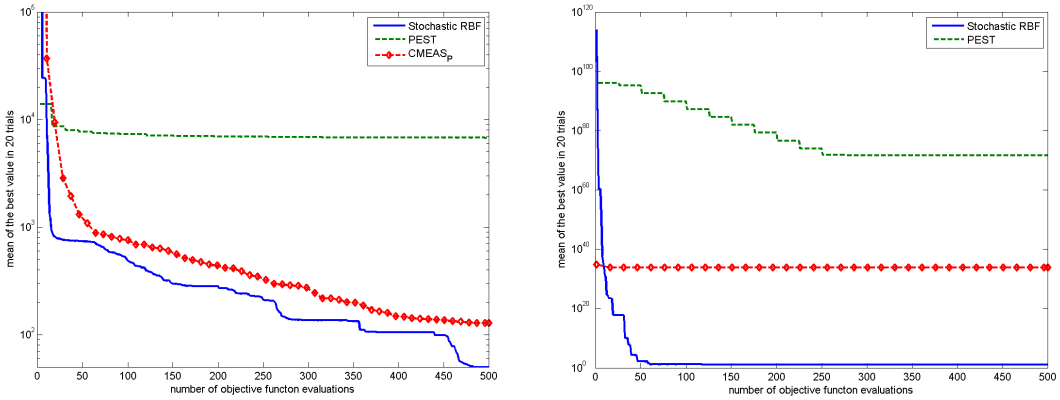

Figure 5.1: Algorithm performance comparisons for test functions calibration in 500 function evaluations: (a) Ackley Function case. (b)Hartmann6 Function case

CHAPTER 6
DISCUSSION AND CONCLUSIONS

In this study, we introduced a novel optimization algorithm for computationally expensive groundwater models calibration and compared this algorithm to two widely used methods. In addition, we proposed a new Hybrid method of combining the new algorithm with one of the earlier methods and compared it to the other methods. All the optimization methods were applied to three cases of 28-parameter calibration of a groundwater model that was based on data for a very large aquifer near Beijing, and two synthetic test function calibration problems. The results of automatic calibration through optimization for all the cases indicate the efficacy of Stochastic RBF for computationally expensive nonlinear optimization problems.

In this study, we compared the new algorithm (Stochastic RBF) with PEST and CMAES_P, and Stochastic RBF gave good results to case studies that had different computational complexity of the underlying optimization problem. Since the goal is to achieve accurate modeling of large groundwater regions with limited computation effort for model calibration, it became evident from this study that Stochastic RBF is more suitable when it is infeasible to perform a relatively large number of computationally expensive model simulations. In terms of efficiency, Stochastic RBF was superior to PEST and CMAES_P in all the cases within both 300 function evaluations and 500 function evaluations (see Figure 4.3 and Figure 4.6). Although it was not as good as the mixed method in the coarse Case, it still performed much better

than the mixed method in both fine Case and hypothetical Case (see Figure 4.7) which have more complexity models than the coarse grid model. This result further illustrates that Stochastic RBF is an effective tool for calibrating parameter for computationally expensive groundwater models. The reliability of Stochastic RBF as indicated by the boxplots as well as values of standard deviation (see Table 4.4, Table 4.5 and Table 4.6) also provided clear evidence of superiority of this algorithm over other methods. Table 4.6 provided the results of 4 measures of goodness of fit both for the original calibration and for an independent set of validation data for each of the three Cases, which in a statistical sense shows that Stochastic RBF has better results in all the measure of goodness of fit criteria.

Stochastic RBF applied in this study can be extended to any optimization problem that requires computationally expensive simulations for cost evaluation within a limited number of function evaluations. In fact, there are numerous computationally expensive nonlinear models in the real world, especially in environmental field, which Stochastic RBF will be suitable to be applied to.

BIBLIOGRAPHY

[1] Abbaspour, K., R. Schulin, and M. van Genuchten (2001), Estimating unsaturated soil hydraulic parameters using ant colony optimization, *Advances in Water Resources*, *24*(8), 827–841.

[2] Ackley, D. H. (1987), A connectionist machine for genetic hillclimbing, Kluwer Academic Pub.

[3] Agyei, E., and K. Hatfield (2006), Enhancing gradient-based parameter estimation with an evolutionary approach, Journal of Hydrology, 316(1-4), 266–280, doi:10.1016/j.jhydrol.2005.05.010.

[4] Aral, M. M. (1985), *Aquifer Parameter Prediction in Leaky Aquifers*, Elsevier Science.

[5] Buhmann, M. D. (2003), *Radial Basis Functions: Theory and Implementations*, Cambridge University Press. [online] Available from: http://www.library.cornell.edu

[6] Carrera, J. S., A. S. Alcolea, A. N. Medina, J. Hidalgo, and L. J. Slooten (2005), Inverse problem in hydrogeology, *Hydrogeol J*, *13*(1), 206–222, doi:10.1007/s10040-004-0404-7.

[7] CHENG, J.-M., and W. W. G. Yeh (1992), A Proposed Quasi-Newton Method for Parameter-Identification in a Flow and Transport-System, *Advances in Water Resources*, *15*(4), 239–249.

[8] Conn, A. R., K. Scheinberg, and L. N. Vicente (2006), Geometry of interpolation sets in derivative free optimization, *Math. Program.*, *111*(1-2), 141–172, doi:10.1007/s10107-006-0073-5.

[9] Dixon, L., and G. Szego (1978), Dixon: The global optimization problem: an introduction - Google Scholar, *Optimization*.

[10] Doherty, J. (2010), PEST: Model-Independent Parameter Estimation, user's manual, 5th ed., Watermark Numerical Computing, Brisbane, Australia.

[11] Doherty, J. (2012), *Addendum to the PEST Manual*, Water Numerical Computing, Brisbane, Australia.

[12] Doherty, J., and D. Welter (2010), A short exploration of structural noise, *Water Resour. Res.*, *46*(5), doi:10.1029/2009WR008377.

[13] Doherty, J., and J. M. Johnston (2003), METHODOLOGIES FOR CALIBRATION AND PREDICTIVE ANALYSIS OF A WATERSHED MODEL1, *JAWRA Journal of the American Water Resources Association*, *39*(2), 251–265.

[14] Duan, Q. (1991), A global optimization strategy for efficient and effective calibration of hydrologic models. - The University of Arizona Campus Repository,

[15] Duan, Q. Y., S. Sorooshian, and V. GUPTA (1994), Optimal Use of the Sce-Ua Global Optimization Method for Calibrating Watershed Models, *Journal of Hydrology*, *158*, 265–284, doi:10.1016/0022-1694(94)90057-4.

[16] Duan, Q. Y., V. K. Gupta, and S. Sorooshian (1993), Shuffled complex evolution approach for effective and efficient global minimization, *J. Optim. Theory Appl.*, *76*(3), 501–521, doi:10.1007/BF00939380.

[17] Finsterle, S. (2007), iTOUGH2 user's guide, *LBNL-40040*, 130.

[18] Finsterle, S., and Y. Zhang (2011), Solving iTOUGH2 simulation and optimization problems using the PEST protocol, *Environmental Modelling and Software*, *26*(7), 959–968, doi:10.1016/j.envsoft.2011.02.008.

[19] Giacobbo, F., M. Marseguerra, and E. Zio (2002), Solving the inverse problem of parameter estimation by genetic algorithms: the case of a groundwater contaminant transport model, *Ann Nucl Energy*, *29*(8), 967–981.

[20] Ginn, T. R., and J. H. Cushman (1990), Inverse methods for subsurface flow: A critical review of stochastic techniques, *Stochastic Hydrology and Hydraulics*, *4*(1), 1–26, doi:10.1007/BF01547729.

[21] Goldberg, D. (1989), *Genetic Algorithms in Search, Optimization, and Machine Learning*, 1st ed., Addison-Wesley Professional.

[22] Hanna, S. K. (1995), An iterative Monte Carlo technique for estimating conditional means and variances of transmissivity and hydraulic head fields., The University of Arizona.

[23] Hansen, N., and A. Ostermeier (2001), Completely derandomized self-adaptation in evolution strategies, *Evolutionary computation*, *9*(2), 159–195.

[24] Hansen, N., and S. Kern (2004), Evaluating the CMA evolution strategy on multimodal test functions, *Parallel Problem Solving from Nature-PPSN VIII*, 282–291.

[25] Heidari, M., and S. Ranjithan (1998), A hybrid optimization approach to thf estimation of distributed parameters in two-dimensional confined aquifers, *J Am Water Resour as*, *34*(4), 909–920.

[26] Hill, M. C., and C. R. Tiedeman (2007), *Effective Groundwater Model Calibration: With Analysis of Data, Sensitivities, Predictions, and Uncertainty*, 1st ed., Wiley-Interscience.

[27] Hill, M., R. Cooley, and D. Pollock (1998), A controlled experiment in ground water flow model calibration, *Ground Water*, *36*(3), 520–535.

[28] Hyun, Y., and K. K. Lee (1998), Model identification criteria for inverse estimation of hydraulic parameters, *Ground Water*, *36*(2), 230–239.

[29] Karahan, H., and M. T. Ayvaz (2005), Groundwater parameter estimation by optimization and dual reciprocity finite differences method, *Journal of Porous Media*, *8*(2).

[30] Karahan, H., and M. T. Ayvaz (2006), Forecasting Aquifer Parameters Using Artificial Neural Networks, *Journal of Porous Media*, *9*(5).

[31] Karahan, H., and M. T. Ayvaz (2008), Simultaneous parameter identification of a heterogeneous aquifer system using artificial neural networks, *Hydrogeol J*, *16*(5), 817–827, doi:10.1007/s10040-008-0279-0.

[32] Kuiper, L. K. (1986), A comparison of several methods for the solution of the inverse problem in two dimensional steady state groundwater flow modelling, *Water Resour. Res.*, *22*(5), 705–714.

[33] LIN, A., and W. YEH (1974), Identification of Parameters in an Inhomogeneous Aquifer by Use of Maximum Principle of Optimal Control and Quasi-Linearization, *Water Resour. Res.*, *10*(4), 829–838, doi:10.1029/WR010i004p00829.

[34] Lingireddy, S. (1998), Aquifer parameter estimation using genetic algorithms and neural networks, *Civil engineering and environmental systems*, *15*(2), 125–144.

[35] Marquardt, D. W. (1963), An Algorithm for Least-Squares Estimation of Nonlinear Parameters, *J Soc Ind Appl Math*, *11*(2), 431–441.

[36] McLaughlin, D., and L. R. Townley (1996), A reassessment of the groundwater inverse problem, *Water Resour. Res.*, *32*(5), 1131–1161.

[37] Mugunthan, P., C. A. Shoemaker, and R. G. Regis (2005), Comparison of function approximation, heuristic, and derivative-based methods for automatic calibration of computationally expensive groundwater bioremediation models, *Water Resour. Res.*, *41*(11), doi:10.1029/2005WR004134.

[38] Nash, J. E., and J. Sutcliffe (1970), River flow forecasting through conceptual models part I—A discussion of principles, *Journal of Hydrology*, *10*(3), 282–290.

[39] OLSTHOORN, T. (1995), Effective Parameter Optimization for Groundwater Model Calibration, *Ground Water*, *33*(1), 42–48.

[40] Poeter, E. P., United States Environmental Protection Agency, International Ground Water Modeling Center.,, and Geological Survey U.S (2005), *UCODE_2005 and six other computer codes for universal sensitivity analysis, calibration, and uncertainty evaluation constructed using the JUPITER API; JUPITER: Joint Universal Parameter Identification and Evaluation of Reliability, API: Application Programming Interface*, U.S. Dept. of the Interior, U.S. Geological Survey, Reston, Va.

[41] Powell, M. (2006), The NEWUOA software for unconstrained optimization without derivatives, *Large-scale nonlinear optimization*, 255–297.

[42] Powell, M. (1992), The Theory of Radial Basis Function Approximation in 1990, in *Advances in Numerical Analysis: Volume II: Wavelets, Subdivision Algorithms, and Radial Basis Functions*, edited by W. Light, pp. 105–210, Oxford University Press, USA.

[43] Powell, M. (1999), Recent research at Cambridge on radial basis functions, *International Series of Numerical Mathematics*, 215–232.

[44] Powell, M. J. D. (2002), UOBYQA: unconstrained optimization by quadratic approximation, *Math. Program.*, *92*(3), 555–582, doi:10.1007/s101070100290.

[45] Powell, M. J. D. (2008), Developments of NEWUOA for minimization without derivatives, *IMA Journal of Numerical Analysis*, *28*(4), 649–664, doi:10.1093/imanum/drm047.

[46] Prasad, K., and A. Rastogi (2001), Estimating net aquifer recharge and zonal hydraulic conductivity values for Mahi Right Bank Canal project area,

India by genetic algorithm, *Journal of Hydrology*, *243*, 149–161, doi:10.1016/S0022-1694(00)00364-4.

[47] Pruess, K., C. Oldenburg, and G. Moridis (1999), TOUGH2 user's guide, version 2.0,

[48] Regis, R. G. (2004), Global optimization of computationally expensive functions using serial and parallel radial basis function algorithms,, 684.

[49] Regis, R. G., and C. A. Shoemaker (2004), Local Function Approximation in Evolutionary Algorithms for the Optimization of Costly Functions, *IEEE Trans. Evol. Computat.*, *8*(5), 490–505, doi:10.1109/TEVC.2004.835247.

[50] Regis, R. G., and C. A. Shoemaker (2007), A Stochastic Radial Basis Function Method for the Global Optimization of Expensive Functions, *INFORMS Journal on Computing*, *19*(4), 497–509, doi:10.1287/ijoc.1060.0182.

[51] Regis, R. G., and C. A. Shoemaker (2005), Constrained Global Optimization of Expensive Black Box Functions Using Radial Basis Functions, *J Global Optim*, *31*(1), 153–171, doi:10.1007/s10898-004-0570-0.

[52] Remson, I., G. M. Hornberger, and F. J. Molz (1971), *Numerical methods in subsurface hydrology with an introduction to the finite element method*, Wiley, New York [etc.].

[53] Shoemaker, C. A., R. G. Regis, and R. C. FLEMING (2007), Watershed calibration using multistart local optimization and evolutionary optimization with radial basis function approximation, *Hydrological Sciences Journal*, *52*(3), 450–465, doi:10.1623/hysj.52.3.450.

[54] Solomatine, D. P., Y. B. Dibike, and N. Kukuric (1999), Automatic calibration of groundwater models using global optimization techniques, *Hydrological Sciences Journal–Journal des Sciences Hydrologiques*, *44*(6), 879–894.

[55] Sun, N. Z. (1994), *Inverse Problems in Groundwater Modeling*, Theory and Applications of Transport in Porous Media, 1st ed., Springer.

[56] Sun, N.-Z., M.-C. Jeng, and W. W. G. Yeh (1995), A proposed geological parameterization method for parameter identification in three-dimensional groundwater modeling, *Water Resour. Res.*, *31*(1), 89–102, doi:10.1029/94WR02276.

[57] Tolson, B. A., and C. A. Shoemaker (2007), Dynamically dimensioned search algorithm for computationally efficient watershed model calibration, *Water Resour. Res.*, *43*(1), doi:10.1029/2005WR004723.

[58] Tung, C.-P., C.-C. Tang, and Y.-P. Lin (2003), Improving groundwater-flow modeling using optimal zoning methods, *Environmental Geology*, *44*(6), 627–638, doi:10.1007/s00254-003-0822-1.

[59] Wang, G., J. Wei, and Y. Zhang (2004), Analysis of the Sustainability of the Development of a Small Phreatic Aquifer in Northern China, *Water International*, *29*(4), 467–474, doi:10.1080/02508060408691809.

[60] Wang, H. F., and M. P. Anderson (1982), *Introduction to Groundwater Modeling: Finite Difference and Finite Element Methods (Series of Books in Geology)*, edited by A. Cox, W H Freeman & Co (Sd).

[61] Wei, J. H. (2001), Groundwater Geographical Information System-Integration, Visualization and Case Study, China.

[62] Wei, J., G. Wang, and C. J. Li (2002), GIS-based groundwater resource evaluation, *J. Tsinghua Univ. (Sci. & Tech)*, 1104–2207.

[63] Wild, S. M., R. G. Regis, and C. A. Shoemaker (2009), ORBIT: Optimization by radial basis function interpolation in trust-regions, *SIAM Journal on Scientific Computing*, *30*(6).

[64] Yeh, W. (1986), Review of Parameter Identification Procedures in Groundwater Hydrology, *Water Resour. Res.*, *22*(2), 95–108.

[65] Zheng, C., and P. Wang (1996), Parameter structure identification using tabu search and simulated annealing, *Advances in Water Resources*, *19*(4), 215–224, doi:10.1243/09544054JEM334.